Posudek práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy

☐ posudek vedoucího ☒ posudek oponenta
☐ bakalářské práce ☒ diplomové práce

Autor/ka: Bc. Ondřej Dušek
Název práce: Analysis of topological magnetic phases using generative machine learning models
Studijní program a obor: Physics of Condensed Matter and Materials
Rok odevzdání: 2023

Jméno a tituly ~~vedoucího~~/oponenta: RNDr. Martin Žonda, Ph.D.
Pracoviště: Department of Condensed Matter Physics, Faculty of Mathematics and Physics, Charles University
Kontaktní e-mail: martin.zonda@matfyz.cuni.cz

Odborná úroveň práce:
☐ vynikající ☒ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Věcné chyby:
☒ téměř žádné ☐ vzhledem k rozsahu přiměřený počet ☐ méně podstatné četné ☐ závažné

Výsledky:
☒ originální ☐ původní i převzaté ☐ netriviální kompilace ☐ citované z literatury ☐ opsané

Rozsah práce:
☐ veliký ☒ standardní ☐ dostatečný ☐ nedostatečný

Grafická, jazyková a formální úroveň:
☐ vynikající ☒ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Tiskové chyby:
☐ téměř žádné ☒ vzhledem k rozsahu a tématu přiměřený počet ☐ četné

Celková úroveň práce:
☐ vynikající ☒ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Slovní vyjádření, komentáře a připomínky ~~vedoucího~~/oponenta:

Bc. Ondřej Dušek in his master thesis utilized modern machine learning methods to investigate skyrmion lattices in two-dimensional magnetic systems. He used and benchmarked several realizations of a special type of network known as variational autoencoder. These networks are trained to reproduce the inputs (spin-configuration) by sequentially downsizing the networks layers down to a bottleneck and then upsizing them to the original data size again. The bottleneck is known as latent space. Despite having much less parameters than the original data, it can be used to reconstruct them. Interestingly, the latent space of 500 parameters used in the work was sufficient to encode complicated skyrmionics configurations on a lattice containing 200x200 classical Heisenberg spins. Moreover, authors PCA analysis of the latent space showed, that it might be even possible to shrink this layer to almost half without losing the networks generative power. In addition, the work analyzes the latent space and what physical features it encodes. The most challenging features in this respect proved to be lattice defects and, therefore, a great focus is dedicated to them in the work. It is shown that the mean square error between the test data and their reconstruction is a good indicator of the number of defects in the samples. Another advantage of a trained generative network is that it can be used to generate new skyrmionics lattice configurations without the need of the resources hungry Monte Carlo simulations. One just needs to understand and utilize the latent space vector. Surprisingly, when tested this led to latices containing many more defects than in the training data. This was attributed to the random sampling of the latent space vector and the fact that without physical constrictions the generation of faulty lattice is more probable than an ideal one.

In my opinion it is scientifically a very good work with interesting results in pair of what is currently published in the field. However, I have several complaints about the presentation. My most serious and general one is what I would call insufficient physical justifications. This is most obvious in the introduction and what is summarized as theoretical part. There I would appreciate a clearly stated physical motivation why to investigate such systems. How relevant are these systems for real materials? What are the problems that the method used in the thesis can help us with? I think this should be addressed in the defense. This problem is, however, not limited to the introduction. The thesis contains a lot of interesting results, some of which I have mentioned above, but rarely is it stressed what is their significance for our understanding of the system. Basically, I was often missing the physical context, which is crucial for a master thesis in physics. For example, why is it important to investigate the correlation of the latent layer with physical properties on which the network is not trained?

My second general complaint is about issues mostly related to figures and their labels. There is almost no visual distinction between the figure labels and the main text and as the figures are large it happens that the label continues to the top of the next page. When reading the main one gets confused by this. There is also strange typesetting where text often ends in the half of the page, so I was expecting a new section, but what follows is just a large figure.

The PCA deserves a more thorough introduction. From these less than 20 lines it is simply not clear how it works.

Nevertheless, despite some shortcomings of the work listed above, it is a solid thesis presenting real research of a timely problem using modern methods.

Případné otázky při obhajobě a náměty do diskuze:

I have also some more technical questions to address in defense:

1. You state that you use T=0.006, how high is this temperature with respect to real systems, i.e., what is typical J.
2. In Figure 3.13 there is a gap in the data around 60 defects. Why is this so? Also, when you have dislocations, how many defects do they represent in your analysis?

3. Could you explain what is plotted in figure 3.15 in more detail and what is the relevance of discussed correlations between Mz, D, Q with principal components for interpretation of the data?
4. If I understand it correctly, you have generated a new configuration by randomly sampling the latent vectors with gaussian distribution around the latent vectors representing the required type of configuration. This led to a surprising number of defects. However, you have already done the PCA analysis. The principal components are arranged in descending order with respect to the amount of variance they encode. Could you not use this to produce better data? I.e., instead of using uniform variance in normal noise, you could apply the noise to the principal components respecting the natural variance of each principal component and only then rotate the data to the latent vector and generate data. Would this lead to an improvement? What about some physical criteria, like overall energy or in work mentioned effective field? Could they be used to improve the generated data and if so then how?

Práci

☒ doporučuji

☐ nedoporučuji

uznat jako diplomovou/bakalářskou.

Navrhuji hodnocení stupněm:

☐ výborně ☒ velmi dobře ☐ dobře ☐ neprospěl/a

Místo, datum a podpis ~~vedoucího~~/oponenta: